## SEQUENCE LISTING

```
<110> KOBAYASHI, KAZUO
      TAKEUCHI, MAKOTO
      IWAMATSU, AKIHIKO
      YAMAMOTO, KENJI
      KUMAGAI, HIDEHIKO
      YOSHIDA, SATOSHI
```

<120> ENDO-BETA-N-ACETYLGLUCOSAMINIDASE GENE

<130> 081356/0154

<140> 09/700,993

<141> 2000-11-22

<150> JP98/141717

<151> 1998-05-22

<160> 37

<170> PatentIn Ver. 2.1

<210> 1

<211> 2369

<212> DNA

<213> Mucor hiemalis

```
<400> 1
gtcgacccac gcgtccgcgg acgcgtgggc ggacgcgtgg gcggacgcgt gggttttatt 60
ttacataaat atgccttcac ttcaattgca acctgatgac aaactagcac ctgtttcttt 120
tgcacttaag tctatgaatg agttgaggga ctggacgcca gacgaaaaga taaagtttaa 180
cgtttcaagc gtggcactac agcctcgtgt gaaaaacgcc ctgaaacctc aattattgtt 240
aactcatgat atggcaggag gatataaaga agataaaaat attcaaggaa acaattataa 300
agacatttat aacattcaat attggcattt agctgatact tttgtatatt tctctcatga 360
gcgagttagc attcctccag tcaattggac aaatgcttgt catagaaatg gtgtaaagtg 420
tttaggtact tttttagtag aaggaaataa ccaaatgcat gaaatggaag ccttgcttca 480
cggtccacct ttacttaata acactgacga ccctatgaga ttatggagtc cgtattatgc 540
agaccaatta gttgctattg ctaaacacta tggttttgat ggctggttgt tcaatattga 600
atgcgaattc tttccttttc ctacaaatcc aaaattcaaa gctgaagagt tggcaaagtt 660
tctacactat tttaaggaaa aattgcataa cgaaatacct ggatctcaac tcatttggta 720
cgacagcatg acaaatgaag gagaaatcca ctggcagaac cagctcacat ggaaaaatga 780
gttatttttt aaaaacacgg atggtatttt tttgaattat tggtggaaaa aagaataccc 840
tgaaatggcg cgtagagtag ctgaaggaat aggtagatca ggtttagaag tttattttgg 900
tacagatgta tggggaaggc atacttatgg tggcggtggt ttcaaatcat ataagggtgt 960
aaaaactgcc tactctgcaa tgacatcttc tgcattattt ggtatggcat ggacatacga 1020
gcatttcgaa aagtctgaat ttgaaaagat ggatcgtttg ttttggtgtg gtggtaaata 1080
ctctgactat cctccccac ctcctaaaaa cccagatgac gaaaaagaag tagaaagcga 1140
tgatagtgaa gatgagctca tgtacggaca caagaaaggt attgctgaca cggtagaatc 1200
tattcctgta ccaggaacag attggtttgt taccaatttt gatagggggt ttggaaatag 1260
gttttattat agaggaaaga gattactttc tcagccttgg tcccatttat cgcatcaagc 1320
tattctcccc aataaaagct atcgaaatcc agagatttat cccactgatc aaaacattaa 1380
aatcactagt tetetegatt gegateatgg agettttett ggtggaacet egettattat 1440
caaaggccaa cgtttcaatc atagagaatc gcatgatgtt gaaactgaaa ttagtatacc 1500
tctgtataag ctttcattag atgctagtaa aggatgctca ttgcgttata tttatagaac 1560
tttgttgatg aaagatgtaa agttgacagt agcatgtcac ttttcgttaa aaacaaacga 1620
ctcagttaat ttcttcaagg tatggcagcc agatgaaaat ttctcttttg aatatgatga 1680
tggaatgaga gccactgtta caactgaaaa ttctaccgaa agcagatgct ttttattacg 1740
```

Part of
Paper No. 13/B

tacaacagaa gaagatacag gagaaaatga ttggataaca aaaactatta atgtgcctgc 1800 tgttccagaa ggaagtcaat tatacattac aagacttgaa gtgagcgtag tattagatac 1860 agctggttta gtaggtcttg ttaatcaagt tattgcttgc ttgggatata ttagcatcat 1920 accaactata aattctggaa taaaaacaga ttcatcacgc attattcagg atctcttttg 1980 gaaagatcag aaatatacca aaatcggaaa agaaagttta gacgacatag ctcaagaaga 2040 agttcataga tattatggaa cattgaactg ggaaaacaca gcaaatgtag taaacgcttg 2100 ggaggaaata gattactaca acgtttttta caaagaaagt gacgactctg caactcgcat 2160 ctttttagga acagcatcct gtaatcaatt tcgtgtatct ggtttagata ttattttatc 2220 taagcatagat ttgtcattaa actaggactt taacaaagaa ggatacatct cttcaagtgg 2280 tagcatagat ttgtcattaa actaggactt gaaataaaaa aaaaaaaaa ggcggccgc 2369

<210> 2 <211> 2235 <212> DNA <213> Mucor hiemalis <220> <221> CDS <222> (1)..(2232) <400> 2 48 atg cct tca ctt caa ttg caa cct gat gac aaa cta gca cct gtt tct Met Pro Ser Leu Gln Leu Gln Pro Asp Asp Lys Leu Ala Pro Val Ser ttt gca ctt aag tct atg aat gag ttg agg gac tgg acg cca gac gaa 96 Phe Ala Leu Lys Ser Met Asn Glu Leu Arg Asp Trp Thr Pro Asp Glu aag ata aag ttt aac gtt tca agc gtg gca cta cag cct cgt gtg aaa 144 Lys Ile Lys Phe Asn Val Ser Ser Val Ala Leu Gln Pro Arg Val Lys 35 aac gcc ctg aaa cct caa tta ttg tta act cat gat atg gca gga gga 192 Asn Ala Leu Lys Pro Gln Leu Leu Thr His Asp Met Ala Gly Gly 50 55 240 tat aaa gaa gat aaa aat att caa gga aac aat tat aaa gac att tat Tyr Lys Glu Asp Lys Asn Ile Gln Gly Asn Asn Tyr Lys Asp Ile Tyr 65 70 75 288 aac att caa tat tgg cat tta gct gat act ttt gta tat ttc tct cat Asn Ile Gln Tyr Trp His Leu Ala Asp Thr Phe Val Tyr Phe Ser His gag cga gtt agc att cct cca gtc aat tgg aca aat gct tgt cat aga 336 Glu Arg Val Ser Ile Pro Pro Val Asn Trp Thr Asn Ala Cys His Arg 100 aat ggt gta aag tgt tta ggt act ttt tta gta gaa gga aat aac caa 384 Asn Gly Val Lys Cys Leu Gly Thr Phe Leu Val Glu Gly Asn Asn Gln 115 120 125 atg cat gaa atg gaa gcc ttg ctt cac ggt cca cct tta ctt aat aac 432 Met His Glu Met Glu Ala Leu Leu His Gly Pro Pro Leu Leu Asn Asn 130 140 135

	_	_		_	_			_	_			_	_	caa Gln		480
														aat Asn 175		528
														gct Ala	_	576
									_	_		_		aac Asn	_	624
									_	_	_			gaa Glu		672
_				_		_								ttt Phe		720
														gaa Glu 255		768
												_		ggt Gly		816
														ggt Gly		864
						_		_			_			gca Ala	_	912
			_	Leu			_	_						ttc Phe	_	960
aag Lys	tct Ser	gaa Glu	ttt Phe	gaa Glu 325	aag Lys	atg Met	gat Asp	cgt Arg	ttg Leu 330	ttt Phe	tgg Trp	tgt Cys	ggt Gly	ggt Gly 335	aaa Lys	1008
														gaa Glu		1056
		_	-	_	_	_	-	_			_			cac His		1104

Î

									•							
									4							
	ggt Gly 370															1152
	ttt Phe					Asp										1200
	gga Gly															1248
_	att Ile					_		_								1296
	caa Gln															1344
	ctt Leu 450				_							_				1392
_	gaa Glu	_		_	_	_		_		_			_		_	1440
	tca Ser															1488
	ttg Leu	_	_		_	_	_	_		_	_	_			_	1536
	aaa Lys			_		_					_		_		_	1584
_	aat Asn 530				-		_	-		_	_	_		_		1632
	gaa Glu				_	_	_	_				_			-	1680
	gat Asp															1728
-	gtt Val		_		_						_		_		_	1776

5 · Ť ·

_	_		_		_			_			gtt Val			_		1824
_	_	_				_					ata Ile 620					1872
		_			_			_	_		ttt Phe			_	_	1920
							_	_		_	gac Asp		_		_	1968
											gaa Glu					2016
_	_		_			_		_			aac Asn	_				2064
_	_	_	_		_		_				gga Gly 700		_		_	2112
			_	_				_			tta Leu		_			2160
_		_		_	_	_			_		tac Tyr					2208
	_		gat Asp 740	_				tag								2235

<210> 3

· 1 ·

<211> 744

<212> PRT

<213> Mucor hiemalis

<400> 3

Met Pro Ser Leu Gln Leu Gln Pro Asp Asp Lys Leu Ala Pro Val Ser 1 5 10 15

Phe Ala Leu Lys Ser Met Asn Glu Leu Arg Asp Trp Thr Pro Asp Glu 20 25 30

Lys Ile Lys Phe Asn Val Ser Ser Val Ala Leu Gln Pro Arg Val Lys 35 40 45

Asn Ala Leu Lys Pro Gln Leu Leu Leu Thr His Asp Met Ala Gly Gly 50 55 60

. 7

Tyr Lys Glu Asp Lys Asn Ile Gln Gly Asn Asn Tyr Lys Asp Ile Tyr 75 Asn Ile Gln Tyr Trp His Leu Ala Asp Thr Phe Val Tyr Phe Ser His 90 Glu Arg Val Ser Ile Pro Pro Val Asn Trp Thr Asn Ala Cys His Arg Asn Gly Val Lys Cys Leu Gly Thr Phe Leu Val Glu Gly Asn Asn Gln 120 Met His Glu Met Glu Ala Leu Leu His Gly Pro Pro Leu Leu Asn Asn Thr Asp Asp Pro Met Arg Leu Trp Ser Pro Tyr Tyr Ala Asp Gln Leu 150 155 Val Ala Ile Ala Lys His Tyr Gly Phe Asp Gly Trp Leu Phe Asn Ile 165 170 Glu Cys Glu Phe Phe Pro Phe Pro Thr Asn Pro Lys Phe Lys Ala Glu 180 185 190 Glu Leu Ala Lys Phe Leu His Tyr Phe Lys Glu Lys Leu His Asn Glu 200 205 Ile Pro Gly Ser Gln Leu Ile Trp Tyr Asp Ser Met Thr Asn Glu Gly Glu Ile His Trp Gln Asn Gln Leu Thr Trp Lys Asn Glu Leu Phe Phe 235 Lys Asn Thr Asp Gly Ile Phe Leu Asn Tyr Trp Trp Lys Lys Glu Tyr 255 Pro Glu Met Ala Arg Arg Val Ala Glu Gly Ile Gly Arg Ser Gly Leu Glu Val Tyr Phe Gly Thr Asp Val Trp Gly Arg His Thr Tyr Gly Gly Gly Gly Phe Lys Ser Tyr Lys Gly Val Lys Thr Ala Tyr Ser Ala Met Thr Ser Ser Ala Leu Phe Gly Met Ala Trp Thr Tyr Glu His Phe Glu 315 Lys Ser Glu Phe Glu Lys Met Asp Arg Leu Phe Trp Cys Gly Gly Lys 330 Tyr Ser Asp Tyr Pro Pro Pro Pro Lys Asn Pro Asp Asp Glu Lys 345 Glu Val Glu Ser Asp Ser Glu Asp Glu Leu Met Tyr Gly His Lys

360

355

Lys Gly Ile Ala Asp Thr Val Glu Ser Ile Pro Val Pro Gly Thr Asp 370 375 380

· 'I ·

Trp Phe Val Thr Asn Phe Asp Arg Gly Phe Gly Asn Arg Phe Tyr Tyr 385 390 395 400

Arg Gly Lys Arg Leu Leu Ser Gln Pro Trp Ser His Leu Ser His Gln 405 410 415

Ala Ile Leu Pro Asn Lys Ser Tyr Arg Asn Pro Glu Ile Tyr Pro Thr 420 425 430

Asp Gln Asn Ile Lys Ile Thr Ser Ser Leu Asp Cys Asp His Gly Ala 435 440 445

Phe Leu Gly Gly Thr Ser Leu Ile Ile Lys Gly Gln Arg Phe Asn His 450 455 460

Arg Glu Ser His Asp Val Glu Thr Glu Ile Ser Ile Pro Leu Tyr Lys 465 470 475 480

Leu Ser Leu Asp Ala Ser Lys Gly Cys Ser Leu Arg Tyr Ile Tyr Arg 485 490 495

Thr Leu Leu Met Lys Asp Val Lys Leu Thr Val Ala Cys His Phe Ser 500 505 510

Leu Lys Thr Asn Asp Ser Val Asn Phe Phe Lys Val Trp Gln Pro Asp 515 520 525

Glu Asn Phe Ser Phe Glu Tyr Asp Asp Gly Met Arg Ala Thr Val Thr 530 535 540

Thr Glu Asn Ser Thr Glu Ser Arg Cys Phe Leu Leu Arg Thr Thr Glu 545 550 555 560

Glu Asp Thr Gly Glu Asn Asp Trp Ile Thr Lys Thr Ile Asn Val Pro 565 570 575

Ala Val Pro Glu Gly Ser Gln Leu Tyr Ile Thr Arg Leu Glu Val Ser 580 585 590

Val Val Leu Asp Thr Ala Gly Leu Val Gly Leu Val Asn Gln Val Ile 595 600 605

Ala Cys Leu Gly Tyr Ile Ser Ile Ile Pro Thr Ile Asn Ser Gly Ile 610 615 620

Lys Thr Asp Ser Ser Arg Ile Ile Gln Asp Leu Phe Trp Lys Asp Gln 625 630 635 640

Lys Tyr Thr Lys Ile Gly Lys Glu Ser Leu Asp Asp Ile Ala Gln Glu 645 650 655

Glu Val His Arg Tyr Tyr Gly Thr Leu Asn Trp Glu Asn Thr Ala Asn
660 665 670

Val Val Asn Ala Trp Glu Glu Ile Asp Tyr Tyr Asn Val Phe Tyr Lys Glu Ser Asp Asp Ser Ala Thr Arg Ile Phe Leu Gly Thr Ala Phe Cys 695 Asn Gln Phe Arg Val Ser Gly Leu Asp Ile Ile Leu Ser Lys Leu Pro 715 Lys Ile Val Ile Glu Ala Val Asn Lys Glu Gly Tyr Ile Ser Ser Ser 725 730 Gly Ser Ile Asp Leu Ser Leu Asn 740 <210> 4 <211> 10 <212> PRT <213> Artificial Sequence <220> <223> Description of Artificial Sequence: Partial amino acid sequence of endo-beta-N-acetylglucosaminidase Pro Ser Leu Gln Leu Gln Pro Asp Asp Lys 5 <210> 5 <211> 20 <212> DNA <213> Artificial Sequence <223> Description of Artificial Sequence: Oligonucleotide based on the partial amino acid sequence of endo-beta-N-acetylglucosaminidase <220> <221> modified\_base <222> (12) <223> a, g, c or t <400> 5 20 carttrcarc cngaygayaa <210> 6

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Partial amino acid sequence of endo-beta-N-acetylglucosaminidase

9

```
<400> 6
Ser Tyr Arg Asn Pro Glu Ile Tyr Pro Thr Asp Gln Asn Ile Lys
<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Oligonucleotide
      based on the partial amino acid sequence
      of endo-beta-N-acetylglucosaminidase
<220>
<221> modified_base
<222> (6)
<223> a, g, c or t
<4.00> 7
cchacngayc araayatyaa
                                                                    20
<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: Oligonucleotide
      based on the partial amino acid sequence
      of endo-beta-N-acetylglucosaminidase
<220>
<221> modified_base
<222> (15)
<223> a, g, c or t
<400> 8
ttratrttyt grtcngtdgg
                                                                   20
<210> 9
<211> 16
<212> PRT
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Partial amino acid
      sequence of endo-beta-N-acetylglucosaminidase
<400> 9
Gly Gln Arg Phe Asn His Arg Glu Ser His Asp Val Glu Thr Glu Ile
                                      10
```

<210> <211> <212> <213>	20	
	Description of Artificial Sequence: Oligonucleotide based on the partial amino acid sequence of endo-beta-N-acetylglucosaminidase	
<400> tgrttr	10 aadc gytgdccytt	20
<210><211><211><212><213>	23	
	Description of Artificial Sequence: Oligonucleotide sequence of 5' terminal region of endo-beta-N-acetylglucosaminidase gene	
<400> atgcct	11 tcac ttcaattgca acc	23
<210><211><212><212><213>	23	
	Description of Artificial Sequence: Oligonucleotide sequence of 3' terminal region of endo-beta-N-acetylglucosaminidase gene	
<400> ctagtt	12 taat gacaaatcta tgc	23
<210><211><211><212><213>	20	
	Description of Artificial Sequence: Oligonucleotide designed based on the sequence of endo-beta-N-acetylglucosaminidase gene	
<400> cactta	13 agtc tatgaatgag	20
<210>		

, ,

```
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: Oligonucleotide
      designed based on the sequence of endo-beta-N-
      acetylglucosaminidase gene
<400> 14
                                                                   20
aatctctgga tttcgatagc
<210> 15
<211> 41
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: Oligonucleotide
      designed based on the sequence of endo-beta-N-
      acetylglucosaminidase gene
<400> 15
ggggcggccg cttttatttt acataaatat gccttcactt c
                                                                   41
<210> 16
<211> 38
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Oligonucleotide
      designed based on the sequence of endo-beta-N-
      acetylglucosaminidase gene
<400> 16
cccgcggccg cctagtttaa tgacaaatct atgctacc
                                                                   38
<210> 17
<211> 10
<212> PRT
<213> Mucor hiemalis
Pro Ser Leu Gln Leu Gln Pro Asp Asp Lys
                  5
<210> 18
<211> 16
<212> PRT
<213> Mucor hiemalis
Lys Ser Tyr Arg Asn Pro Glu Ile Tyr Pro Thr Asp Gln Asn Ile Lys
                 5
                                     10
                                                          15
```

```
<210> 19
<211> 14
<212> PRT
<213> Mucor hiemalis
<400> 19
Lys Phe Asn Val Ser Ser Val Ala Leu Gln Pro Arg Val Lys
<210> 20
<211> 11
<212> PRT
<213> Mucor hiemalis
<220>
<221> MOD_RES
<222> (2)
<223> Met or Ser
<400> 20
Lys Xaa Asp Arg Leu Phe Leu Cys Gly Gly Lys
                 5
<210> 21
<211> 17
<212> PRT
<213> Mucor hiemalis
<220>
<221> MOD RES
<222> (2)
<223> Gly or Met
<220>
<221> MOD_RES
<222> (3)
<223> Gln or Ala
<220>
<221> MOD_RES
<222> (4)
<223> Arg or Leu
<220>
<221> MOD RES
<222> (6)
<223> Asn or Pro
<220>
<221> MOD RES
<222> (8)
<223> Arg or Leu
<220>
```

.

```
<221> MOD RES
<222> (9)
<223> Glu or Leu
<220>
<221> MOD RES
<222> (10)
<223> Ser or Leu
<220>
<221> MOD_RES
<222> (11)
<223> His or Thr
<400> 21
Lys Xaa Xaa Xaa Phe Xaa His Xaa Xaa Xaa Asp Val Glu Thr Glu
                                     10
Ile
<210> 22
<211> 16
<212> PRT
<213> Mucor hiemalis
<400> 22
Lys Glu Gly Tyr Ile Ser Ser Ser Gly Ser Ile Asp Leu Ser Leu Asn
<210> 23
<211> 9
<212> PRT
<213> Mucor hiemalis
<400> 23
Lys Asn Ile Gln Gly Asn Asn Tyr Lys
<210> 24
<211> 11
<212> PRT
<213> Mucor hiemalis
<400> 24
Lys Tyr Ser Asp Tyr Pro Pro Pro Pro Pro Lys
<210> 25
<211> 8
<212> PRT
<213> Mucor hiemalis
```

```
<400> 25
Lys Leu Ser Leu Asp Ala Ser Lys
<210> 26
<211> 13
<212> PRT
<213> Mucor hiemalis
<400> 26
Lys Asn Thr Asp Gly Ile Phe Leu Asn Tyr Trp Trp Lys
                                     10
<210> 27
<211> 15
<212> PRT
<213> Mucor hiemalis
<220>
<221> MOD RES
<222> (3)
<223> Carboxymethylcystein
Lys Gly Cys Ser Leu Arg Tyr Ile Tyr Arg Thr Leu Leu Met Lys
                                     10
<210> 28
<211> 7
<212> PRT
<213> Mucor hiemalis
<220>
<221> MOD_RES
<222> (6)
<223> Carboxymethylcysteine
<400> 28
Lys Leu Thr Val Ala Cys His
<210> 29
<211> 15
<212> PRT
<213> Mucor hiemalis
<400> 29
Lys Pro Gln Leu Leu Thr His Asp Met Ala Gly Gly Tyr Lys
<210> 30
<211> 14
```

٠ .

<212> PRT

```
<213> Mucor hiemalis
<400> 30
Lys Ser Met Asn Glu Leu Arg Asp Trp Thr Pro Asp Glu Lys
                5
<210> 31
<211> 10
<212> PRT
<213> Mucor hiemalis
<400> 31
Lys Leu Ala Pro Val Ser Phe Ala Leu Lys
                 5
<210> 32
<211> 23
<212> PRT
<213> Mucor hiemalis
<400> 32
Lys Gly Gln Arg Phe Asn His Arg Glu Ser His Asp Val Glu Thr Glu
1
                 5
                                     10
Ile Ser Ile Pro Leu Tyr Lys
             20
<210> 33
<211> 22
<212> PRT
<213> Mucor hiemalis
<220>
<221> MOD_RES
<222> (8)
<223> Carboxymethylcystein
<400> 33
Lys Ile Thr Ser Ser Leu Asp Cys Asp His Gly Ala Phe Leu Gly Gly
Thr Ser Leu Ile Ile Lys
           20
<210> 34
<211> 19
<212> PRT
<213> Mucor hiemalis
Lys Asn Glu Leu Phe Phe Lys Asn Thr Asp Gly Ile Phe Leu Asn Tyr
Trp Trp Lys
```

```
<210> 35
<211> 9
<212> PRT
<213> Mucor hiemalis
<400> 35
Lys Ile Val Ile Glu Ala Val Asn Lys
<210> 36
<211> 11
<212> PRT
<213> Mucor hiemalis
<400> 36
Ser Ser Arg Ile Ile Gln Asp Leu Phe Trp Lys
                 5
<210> 37
<211> 14
<212> PRT
<213> Mucor hiemalis
<400> 37
Lys Thr Asp Ser Ser Arg Ile Ile Gln Asp Leu Phe Trp Lys
```